# Certificates of positivity in the simplicial Bernstein basis.

Richard Leroy

# Certificates of positivity in the simplicial Bernstein basis

Richard Leroy

**Abstract**

We study in the paper the positivity of real multivariate polynomials over a non-degenerate simplex $V$. We aim at obtaining certificates of positivity, *i.e.* algebraic identities certifying the positivity of a given polynomial on $V$, thus generalizing the work in [BCR]. In order to do so, we use the Bernstein polynomials, which are more suitable than the usual monomial basis.

## Contents



## 1 Introduction

We first introduce some necessary material about the simplicial Bernstein basis, and then introduce the notion of certificate of positivity.

**Multivariate polynomials in the simplicial Bernstein basis**

We first recall the definition of a simplex:

**Definition 1.1.** *Let* $\mathbf{v}_0, \dots, \mathbf{v}_k$ *be* $k+1$ *points of* $\mathbb{R}^k$ *(*$k \geq 1$*).*

*The ordered list* $V = [\mathbf{v}_0, \dots, \mathbf{v}_k]$ *is called simplex of vertices* $\mathbf{v}_0, \dots, \mathbf{v}_k$*.*

*The realization* $|V|$ *of the simplex* $V$ *is the set of* $\mathbb{R}^k$ *defined as the convex hull of the points* $\mathbf{v}_0, \dots, \mathbf{v}_k$*.*

*If the points* $\mathbf{v}_0, \dots, \mathbf{v}_k$ *are affinely independent, the simplex* $V$ *is said to be non-degenerate.*

**Notation 1.2.** *Throughout the paper* $V = [\mathbf{v}_0, \dots, \mathbf{v}_k]$ *will denote a non-degenerate simplex of* $\mathbb{R}^k$*, or, by abuse of notation, its realization* $|V|$*.*

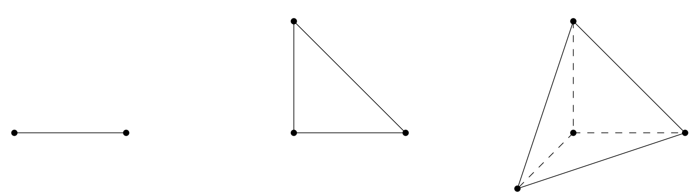
Let $\lambda_0, \dots, \lambda_k$ be the associated barycentric coordinates to $V$, *i.e.* the linear polynomials of $\mathbb{R}[\mathbf{X}] = \mathbb{R}[X_1, \dots, X_k]$ such that

$$\sum_{i=0}^{k} \lambda_i(\mathbf{X}) = 1 \qquad \text{and} \qquad \forall \mathbf{x} \in \mathbb{R}^k,\ \mathbf{x} = \lambda_0(\mathbf{x})\mathbf{v}_0 + \cdots + \lambda_k(\mathbf{x})\mathbf{v}_k.$$

Recall that $V$ is characterized by its barycentric coordinates as follows:

$$V = \bigcap_{i=0}^{k} \{\mathbf{x} \in \mathbb{R}^k | \lambda_i(\mathbf{x}) \geq 0\}.$$

**Example 1.3.** *Let* $(\mathbf{e}_1, \dots, \mathbf{e}_k)$ *denote the canonical basis of* $\mathbb{R}^k$*, and* $\mathbf{0} = (0, \dots, 0)$ *the origin. The simplex* $\Delta = [\mathbf{0}, \mathbf{e}_1, \dots, \mathbf{e}_k]$ *is called standard simplex of* $\mathbb{R}^k$*.*



| | | |
|---|---|---|
| $1 - X_1 \geq 0$ | $1 - X_1 - X_2 \geq 0$ | $1 - X_1 - X_2 - X_3 \geq 0$ |
| $X_1 \geq 0$ | $X_1, X_2 \geq 0$ | $X_1, X_2, X_3 \geq 0$ |

Figure 1: Standard simplices and associated barycentric coordinates in dimension $1, 2, 3$.

The following notation will be useful afterwards:

**Notation 1.4.** *For every multi-index* $\alpha = (\alpha_0, \ldots, \alpha_k) \in \mathbb{N}^{k+1}$*, we write* $|\alpha| = \alpha_0 + \cdots + \alpha_k$*.*

The Bernstein polynomials are defined as follows:

**Definition 1.5.** *Let* $d$ *be a natural number. The Bernstein polynomials of degree* $d$ *with respect to* $V$ *are the polynomials* $\left(B^d_\alpha\right)_{|\alpha|=d}$*, where:*

$$B^d_\alpha = \binom{d}{\alpha}\lambda^\alpha = \frac{d!}{\alpha_0! \ldots \alpha_k!}\prod_{i=0}^{k}\lambda_i^{\alpha_i} \in \mathbb{R}\left[\mathbf{X}\right].$$

Note that the Bernstein polynomials of degree $d$ w.r.t. $V$ take nonnegative values on $V$, and sum up to 1:

$$1 = 1^d = \big(\sum_{i=0}^{k}\lambda_i(\mathbf{X})\big)^d = \sum_{|\alpha|=d}\binom{d}{\alpha}\prod_{i=0}^{k}\lambda_i(\mathbf{X})^{\alpha_i} = \sum_{|\alpha|=d} B^d_\alpha(\mathbf{X}).$$

It is also classical that they form a basis of the vector-space of the polynomials of degree $\leq d$. Thus, every polynomial $f$ of degree $\leq d$ can be uniquely written as

$$f = \sum_{|\alpha|=d} b_\alpha(f, d, V) B^d_\alpha,$$

and the numbers $b_\alpha(f, d, V)$ are called Bernstein coefficients of $f$ of degree $d$ with respect to $V$. We denote by $b(f, d, V)$ the list of all the Bernstein coefficients $b_\alpha(f, d, V)$ (in any order).

### Certificate of positivity

If $b(f, d, V)$ is the list of the Bernstein coefficients (of degree $d$, with respect to $V$) of a polynomial $f$, we define $\mathrm{Cert}(b(f, d, V))$ by:

$$\mathrm{Cert}\left(b(f, d, V)\right): \qquad \begin{cases} \forall |\alpha| = d,\ b_\alpha(f, d, V) \geq 0 \\ \forall i \in \{0, \ldots, k\},\ b_{d\mathbf{e}_i}(f, d, V) > 0, \end{cases}$$

where $(\mathbf{e}_0, \ldots, \mathbf{e}_k)$ denotes the standard basis of $\mathbb{R}^{k+1}$. $\mathrm{Cert}\left(b(f, d, V)\right)$ clearly implies that $f$ is positive on $V$, and then the expression of $f$ in the Bernstein basis of degree $d$ w.r.t. $V$ provides a certificate of positivity for $f$ on $V$, *i.e.* a description of $f$ making obvious that it is positive on $V$.

**Example 1.6.** *Following [BCR], let* $f = 5X^2 - 4X + 1$*. Then* $f$ *is positive on* $[-1, 1]$*, but* $b(f, 2, [-1, 1]) = (10, -4, 2)$*, so that* $\mathrm{Cert}\left(b(f, 2, [-1, 1])\right)$ *does not hold.*
*However, as shown in [BCR],* $\mathrm{Cert}\left(b(f, 21, [-1, 1])\right)$ *holds, providing a certificate of positivity for* $f$ *on* $[-1, 1]$ *with* $22$ *coefficients.*
*A shorter one can be obtained by subdivision, noting that* $\mathrm{Cert}\left(b(f, 2, [-1, 0])\right)$*,* $\mathrm{Cert}\left(b(f, 2, [0, 1/2])\right)$ *and* $\mathrm{Cert}\left(b(f, 2, [1/2, 1])\right)$ *hold.*

In the last example, two techniques were used in order to obtain certificates of positivity: a degree elevation (the interval remain unchanged but the polynomial is seen in a higher degree) and a subdivision (the degree is fixed, but the interval is subdivided). As we will see later on, both of them can be extended to the multivariate case. The idea for deciding if a polynomial is positive (and giving a certificate of positivity) can be sketched as follows:

1. Either by degree elevation or subdivision, the Bernstein coefficients of a polynomial $f$ converge to its graph.
2. If $f$ is positive, then the Bernstein coefficients will eventually satisfy the certificate of positivity (of a higher degree or on each subsimplex).
3. The number of degree elevations / subdivisions can be bounded in terms of the degree, the dimension, the scale of the original Bernstein coefficients and the minimum of $f$.
4. If $f$ is not positive, a stopping criterion is then derived from an explicit lower bound on the minimum of a positive polynomial.

Our plan for this paper is as follows. Section 2 introduces the necessary material about the so-called standard triangulation, which will be crucial in order to study in Section 3 the approximation of a polynomial by its Bernstein coefficients. Based on this latter property, convergence of the control points to the graph of a polynomial is studied in Section 4, the rate being explicitly bounded from above. Based on the results of Sections 4, certificates of positivity are obtained in the multivariate framework in Section 5.

# 2 Standard triangulation

We now introduce the notion of the standard triangulation of an arbitrary simplex of $\mathbb{R}^k$.

## 2.1 Kuhn's triangulation

We first recall the definition of Kuhn's triangulation of the unit cube, which will be useful consequently.

Consider the unit cube $\mathbf{C}_k = [0,1]^k$. If $\sigma \in \mathfrak{S}_k$ is a permutation, let $V^\sigma$ denote the following simplex:

$$V^\sigma = \left\{(x_1,\dots,x_k) \in \mathbb{R}^k \middle|\, 0 \le x_{\sigma(k)} \le \cdots \le x_{\sigma(1)} \le 1\right\}. \tag{2.1}$$

The following result is well-known ([AG],[DM]), though a rigorous proof is hard to find in literature:

**Theorem 2.1.** *The collection of simplices $(V^\sigma)_{\sigma\in\mathfrak{S}_k}$ forms a triangulation of the unit cube, called Kuhn's triangulation, and denoted by $\mathcal{K}(\mathbf{C}_k)$.*

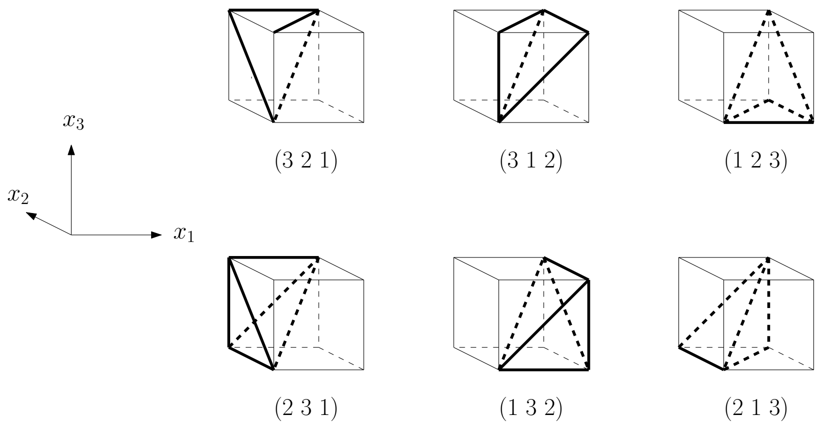



Figure 2: Kuhn's triangulation in dimension 3.

*Proof.*

- It is obvious, from the definition of $V^\sigma$, that $\mathbf{C}_k = \cup\,(V^\sigma)_{\sigma\in\mathfrak{S}_k}$.
- It now suffices to show that the intersection of two simplices $V^\sigma$ and $V^\tau$ is a common face of both simplices.
  Consider two simplices $V^\sigma$ and $V^\tau$, with $\sigma,\tau \in \mathfrak{S}_k$. Let $\mathcal{V}$ be the set of their common vertices. Note that $\mathcal{V}$ contains the vertices $\mathbf{0} = (0,\dots,0)$ and $\mathbf{1} = (1,\dots,1)$. Let $\mathcal{F}$ be the convex hull of $\mathcal{V}$. It is a common face of $V^\sigma$ and $V^\tau$. We now show that $\mathcal{F}$ is in fact the intersection $V^\sigma \cap V^\tau$.
  - Obviously, we have $\mathcal{F} \subset V^\sigma \cap V^\tau$.
  - Consider a point $\mathbf{x} \in V^\sigma \cap V^\tau$. We now show that $\mathbf{x}$ is a convex combination of the vertices of $\mathcal{V}$.
    Since $\mathbf{x} \in V^\sigma$, there exist integers $0 < i_1 < \cdots < i_s$ ($s \geq 1$) such that
    $$\begin{aligned} x_{\sigma(1)} = \cdots = x_{\sigma(i_1)} &> x_{\sigma(i_1+1)} = \cdots = x_{\sigma(i_2)} \\ &> \dots \\ &> x_{\sigma(i_{s-1}+1)} = \cdots = x_{\sigma(i_s)}. \end{aligned}$$
    Since $\mathbf{x} \in V^\sigma$, it also satisfies
    $$\begin{aligned} x_{\tau(1)} = \cdots = x_{\tau(i_1)} &> x_{\tau(i_1+1)} = \cdots = x_{\tau(i_2)} \\ &> \dots \\ &> x_{\tau(i_{s-1}+1)} = \cdots = x_{\tau(i_s)}, \end{aligned}$$
    which easily implies that for all $\ell \in \{1,\dots,s\}$, $\{\sigma(1),\dots,\sigma(i_\ell)\} = \{\tau(1),\dots,\tau(i_\ell)\}$. In particular,
    $$\forall \ell \in \{1,\dots,s\},\ \mathbf{v}^\sigma_{i_\ell} = \mathbf{v}^\tau_{i_\ell} \in \mathcal{V}.$$

Hence $\mathbf{x}$ can be written in the following form:

$$\begin{aligned}\mathbf{x} &= x_{\sigma(i_1)}\mathbf{v}^\sigma_{i_1} + x_{\sigma(i_2)}\left[\mathbf{v}^\sigma_{i_2} - \mathbf{v}^\sigma_{i_1}\right] + \cdots + x_{\sigma(i_s)}\left[\mathbf{v}^\sigma_{i_s} - \mathbf{v}^\sigma_{i_{s-1}}\right] \\ &= \left[1 - x_{\sigma(i_1)}\right]\mathbf{0} + x_{\sigma(i_s)}\mathbf{v}^\sigma_{i_s} + \sum_{\ell=1}^{s-1}\left[x_{\sigma(i_\ell)} - x_{\sigma(i_{\ell+1})}\right]\mathbf{v}^\sigma_{i_\ell},\end{aligned}$$

thus expressing $x$ as a convex combination of vertices in $\mathcal{V}$.

Hence, $\mathcal{F} = V^\sigma \cap V^\tau$, as claimed. □

Kuhn's triangulation can be defined for any affine transformation of the unit cube as follows:

**Proposition 2.2.** *Let $f$ be an affine transformation, and $\mathcal{K}\big(f(\mathbf{C}_k)\big)$ denote the collection:*

$$\mathcal{K}\big(f(\mathbf{C}_k)\big) = \big(f(V^\sigma)\big)_{\sigma\in\mathfrak{S}_k}.$$

*Then $\mathcal{K}\big(f(\mathbf{C}_k)\big)$ is a triangulation, called Kuhn's triangulation, of $f(\mathbf{C}_k)$.*

*Proof.* Obvious. □

An important property of Kuhn's triangulation is its compatibility with translation ([DM]). We first show the restricting Kuhn's triangulation to a face generated by consecutive vertices induces Kuhn's triangulation (in dimension $k-1$) of the face.

**Proposition 2.3.** *The restriction of $\mathcal{K}\big(\mathbf{C}_k\big)$ to a $r-$dimensional face $F$ of $\mathbf{C}_k$ induces Kuhn's triangulation $\mathcal{K}\big(F\big)$ on $F$.*

*Proof.* Without loss of generality, we can assume that $r = k-1$.

- First case : $F = \mathbf{C}_k \cap \{x \in \mathbb{R}^k | x_i = 1\}$ for some $1 \le i \le k$.

  Let $V^\sigma$ be a simplex of $\mathcal{K}(\mathbf{C}_k)$. Then

$$\begin{aligned}&V^\sigma \cap \{x \in \mathbb{R}^k | x_i = 1\} \\ &= \{x \in \mathbb{R}^k |\ x_i = 1 \text{ and } 0 \le x_{\sigma(k)} \le \cdots \le x_{\sigma(1)} \le 1\} \\ &= \left\{x \in \mathbb{R}^k \middle| \begin{array}{c} x_{\sigma(\sigma^{-1}(i))} = \cdots = x_{\sigma(1)} = 1 \\ 0 \le x_{\sigma(k)} \le \cdots \le x_{\sigma(\sigma^{-1}(i)-1)} \le 1 \end{array}\right\}\end{aligned}$$

  is a simplex of dimension at most $k - \sigma^{-1}(i)$ of $\mathbb{R}^k$. Among those simplices, those of dimension $k-1$ are exactly the simplices $V^\sigma \cap \{x \in \mathbb{R}^k | x_i = 1\}$ with $\sigma(1) = i$.

  We now define a one-to-one correspondance between the permutations $\{\sigma \in \mathfrak{S}_k |\ \sigma(1) = i\}$ and the permutations of $\mathfrak{S}_{k-1}$. To every $\sigma \in \mathfrak{S}_k$

satisfying $\sigma(1) = i$, we associate the permutation $\hat{\sigma} \in \mathfrak{S}_{k-1}$ defined as follows:

$$\forall \ell \in \{1, \ldots, k-1\},\ \hat{\sigma}(\ell) = \begin{cases} \sigma(\ell+1) \text{ if } \sigma(\ell+1) < i \\ \sigma(\ell+1) - 1 \text{ else.} \end{cases}$$

Define an affine transformation $\iota : \mathbf{C}_{k-1} \to F$ by

$$\iota(x_1, \ldots, x_{k-1}) = (x_1, \ldots, x_{i-1}, 1, x_i, \ldots, x_{k-1}).$$

Clearly, $F = \iota\big(V^{\hat{\sigma}}\big)$. Hence,

$$\mathcal{K}(F) = \mathcal{K}(\iota(\mathbf{C}_{k-1})) = \big(\iota(V^{\hat{\sigma}})\big)_{\hat{\sigma} \in \mathfrak{S}_{k-1}} = (V^{\sigma} \cap F)_{\sigma \in \mathfrak{S}_k, \sigma(1) = i}.$$

In other words, $\mathcal{K}(F)$ is induced by the restriction of $\mathcal{K}(\mathbf{C}_k)$ to $F$.

- The case $F = \mathbf{C}_k \cap \{x \in \mathbb{R}^k | x_i = 0\}$ for some $1 \leq i \leq k$ is similar. □

**Corollary 2.4.** *Kuhn's triangulation is compatible with translation, i.e. for all $v \in \{0,1\}^k$, $\mathcal{K}(\mathbf{C}_k) \cup \mathcal{K}(v + \mathbf{C}_k)$ is a triangulation of $C \cup (v + \mathbf{C}_k)$.*

*Proof.* It is sufficient to consider the cubes $\mathbf{C}_k$ and $\mathbf{e}_i + \mathbf{C}_k$ for an arbitrary $i \in \{1, \ldots, k\}$, where $(\mathbf{e}_1, \ldots, \mathbf{e}_k)$ is the standard basis of $\mathbb{R}^k$.

Let $F = \mathbf{C}_k \cap \{x \in \mathbb{R}^k | x_i = 1\}$. It is sufficient to show that both Kuhn's triangulations $\mathcal{K}(\mathbf{C}_k)$ and $\mathcal{K}(\mathbf{e}_i + \mathbf{C}_k)$ induce Kuhn's triangulation (in dimension $k-1$) on $F$, which is obvious from the last proposition. □

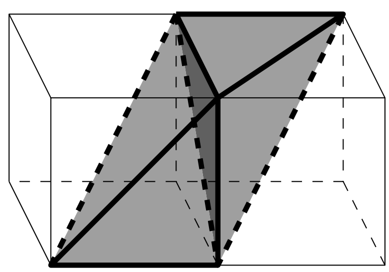


Figure 3: Compatibility of Kuhn's triangulation with translation

## 2.2 Standard triangulation

Based on Kuhn's triangulation, we now present a triangulation of an arbitrary simplex $V$ having nice adjacency and subdivision properties, mentionned as "standard triangulation" in [GP].

### 2.2.1 Particular case

First, consider the simplex $\Lambda = mV^{\mathrm{Id}}$, where Id denotes the identical permutation of $\mathfrak{S}_k$ and $m \geq 2$ is an integer. We aim at triangulating $\Lambda$ into $m^k$ simplices.

Let $F$ be a map from $\{1, \dots, k\}$ to $\{1, \dots, m\}$. Reorder the images of $F$ as follows: $f_1 \leq \dots \leq f_k$, with the convention $f_{k+1} = m$ and $f_0 = 0$.

We now define the point $\mathbf{v}_0^F$ associated to $F$:

$$\mathbf{v}_0^F = (m - f_1, \dots, m - f_k).$$

We also define an application $\sigma_F$ from $\{1, \dots, k\}$ to itself by:

$$\sigma_F(i) = \#\big\{\ell \in \{1, \dots, k\} \big| F(\ell) < F(i)\big\} + \#\big\{\ell \in \{1, \dots, j\} \big| F(\ell) = F(i)\big\}.$$

**Lemma 2.5.** *$\sigma_F$ is a permutation of $\{1, \dots, k\}$.*

*Proof.* It is sufficient to show that $\sigma_F$ is injective. Assume that $i \neq j$, *e.g.* $i < j$.

- First case: $F(i) \neq F(j)$, *e.g.* $F(i) < F(j)$. Then:

$$\begin{aligned}\sigma_F(i) &= \#\big\{\ell \in \{1, \dots, k\} \mid F(\ell) < F(i)\big\} \ + \ \#\big\{\ell \in \{1, \dots, i\} \mid F(\ell) = F(i)\big\} \\ &\leq \#\big\{\ell \in \{1, \dots, k\} \mid F(\ell) < F(i)\big\} \ + \ \#\big\{\ell \in \{1, \dots, k\} \mid F(\ell) = F(i)\big\} \\ &\leq \#\big\{\ell \in \{1, \dots, k\} \mid F(\ell) \leq F(i)\big\} \\ &< \#\big\{\ell \in \{1, \dots, k\} \mid F(\ell) < F(j)\big\} + 1 \leq \sigma_F(j).\end{aligned}$$

  Second case: $F(i) = F(j)$. Then:

$$\begin{aligned}\sigma_F(i) &= \#\big\{\ell \in \{1, \dots, k\} \mid F(\ell) < F(i)\big\} \ + \ \#\big\{\ell \in \{1, \dots, i\} \mid F(\ell) = F(i)\big\} \\ &= \#\big\{\ell \in \{1, \dots, k\} \mid F(\ell) < F(j)\big\} \ + \ \#\big\{\ell \in \{1, \dots, i\} \mid F(\ell) = F(j)\big\} \\ &< \underbrace{\#\big\{\ell \in \{1, \dots, k\} \mid F(\ell) < F(j)\big\} \ + \ \#\big\{\ell \in \{1, \dots, j\} \mid F(\ell) = F(j)\big\}}_{=\sigma_F(j)}.\end{aligned}$$

  In both cases, we have shown that $\sigma_F(i) \neq \sigma_F(j)$.

- The case $j > i$ is similar. ◻

**Remark 2.6.** *Note that, by definition of $\sigma_F$, we have $F(j) = f_{\sigma_F(j)}$ for all $j$. In other words, the permutation $\sigma_F$ sorts the images of $F$ in increasing order.*

The simplex associated to $F$ can now be defined as follows:

$$V^F = \mathbf{v}_0^F + V^{\sigma_F},$$

where $V^{\sigma_F}$ is the simplex associated to the permutation $\sigma_F$ in the Kuhn's triangulation of the unit cube (see Equation (2.1)).

**Remark 2.7.** *The barycentric coordinates $\lambda_i^F$ of the vertices $\mathbf{v}_i^F$ of $V^F$ (with respect to $V$) satisfy:*

$$m\lambda_0^F = (\dots, f_{\ell+1} - f_\ell, \dots), \quad \ell \in \{0, \dots, k\} \tag{2.2a}$$

$$m\lambda_j^F = m\lambda_{j-1}^F + \mathbf{e}_{\sigma_F(j)} - \mathbf{e}_{\sigma_F(j)-1}, \tag{2.2b}$$

*where $(\mathbf{e}_0, \dots, \mathbf{e}_k)$ is the standard basis of $\mathbb{R}^{k+1}$.*

**Remark 2.8.** *If a simplex $V \in T_m(V)$ is given by the unsorted list of its vertices, it is possible to recover the order on its vertices as follows:*

- *The first vertex is the first point in the lexicographic order on the (Euclidean) coordinates.*
- *The successor of a vertex is the only remaining point at Hamming distance 1 from it (recall that the Hamming distance between two points $x, y \in \mathbb{R}^k$ is the number of nonzero coordinates of $x - y$).*

*Consequently, it is also possible to recover the permutation $\sigma_F$.*

*Furthermore, according to Remark 2.7, the values $f_1, \dots, f_k$ can be recovered from the barycentric coordinates of the first vertex of $V$.*

*Finally, knowing $f_1, \dots, f_k$ and $\sigma_F$ enables us to recover the map $F$.*

As expected, we have:

**Theorem 2.9.** *The collection $(V^F)_{F \in \{1,\dots,m\}^{\{1,\dots,k\}}}$ is a triangulation of $\Lambda$.*

*Proof.*

- We first show that $\cup V^F \subset \Lambda$. Consider a simplex $V^F$. Let $\mathbf{x} \in V^F$. Then $\mathbf{x}$ can be written
$$\mathbf{x} = \mathbf{v}_0^F + y, \ y \in V^{\sigma_F}.$$
We now show that $\mathbf{x} \in \Lambda$. It is equivalent to show that for all $i$, $x_{\sigma_F(i)} \geq x_{\sigma_F(i)+1}$. Note that
$$x_{\sigma_F(i)} - x_{\sigma_F(i)+1} = \underbrace{y_{\sigma_F(i)} - y_{\sigma_F(i)+1}}_{\in[-1,1]} + \underbrace{f_{\sigma_F(i)+1} - f_{\sigma_F(i)}}_{\in\mathbb{N}}.$$
If $f_{\sigma_F(i)+1} - f_{\sigma_F(i)} > 0$, then the result is clear.

  Assume that $f_{\sigma_F(i)+1} - f_{\sigma_F(i)} = 0$.
Let $j = \sigma_F^{-1}(\sigma_F(i) + 1)$. Note that $F(j) = f_{\sigma_F(j)} = f_{\sigma_F(i)+1} = f_{\sigma_F(i)} = F(i)$. Hence the definition of $\sigma_F(i)$ and $\sigma_F(j)$ easily implies that $j = i+1$. But then $y_{\sigma_F(i)} - y_{\sigma_F(i)+1} = y_{\sigma_F(i)} - y_{\sigma_F(i+1)} \geq 0$ by definition of $V^{\sigma_F}$, from which the result obviously follows. Hence $\cup V^F \subset \Lambda$.

- We now show that $\Lambda \subset \cup V^F$. Let $\mathbf{x}$ be an interior point of $\Lambda$. Denote by $\lfloor \mathbf{x} \rfloor$ the point $(\lfloor x_1 \rfloor, \ldots, \lfloor x_k \rfloor)$. Define the integers $f_i = m - \lfloor x_i \rfloor$. Note that $1 \leq f_1 \leq \cdots \leq f_k \leq m$. As $\mathbf{x} - \lfloor \mathbf{x} \rfloor \in [0,1]^k$, there exists a permutation $\sigma$ such that $\mathbf{x} - \lfloor \mathbf{x} \rfloor \in V^\sigma$. We now define an application $F : \{1, \ldots, k\} \to \{1, \ldots, d\}$ by

$$F(j) = f_{\sigma(j)}.$$

  Clearly, $\mathbf{x} \in \lfloor \mathbf{x} \rfloor + V^\sigma = \mathbf{v}_0^F + V^\sigma = V^F$. Hence, the interior of $\Lambda$ is contained in $\cup V^F$, which is a closed set. Then $\Lambda \subset \cup V^F$, as $\Lambda$ is the adherence of its own interior.

- It remains to note that $\cup V^F$ is a triangulation, which easily follows from Proposition 2.4. ☐

**Definition 2.10.** *The collection $(V^F)_{F \in \{1,\ldots,m\}^{\{1,\ldots,k\}}}$ is called standard triangulation of degree $m$ of $\Lambda$, and denoted by $T_m(\Lambda)$.*

### 2.2.2 General case

Let $V = [\,\mathbf{v}_0, \ldots, \mathbf{v}_k]$ be a nondegenerate simplex of $\mathbb{R}^k$. It is easy to define its standard triangulation of degree $m \geq 2$ by affine transformation:

**Definition 2.11.** *Let $h : \mathbb{R}^k \to \mathbb{R}^k$ be an affine transformation mapping the simplex $\Lambda = mV^{\mathrm{Id}}$ to the simplex $V$.*
*The standard triangulation of degree $m$ of $V$ is the collection $(h(U))_{U \in T_m(\Lambda)}$, and is denoted $T_m(V)$.*

It is indeed a triangulation of $V$.

Since the barycentric coordinates are preserved by affine transformations, the vertices of the simplices of $T_m(V)$ satisfy the equalities (2.2) of Remark 2.7.

**Adjacencies** The standard triangulation has a nice adjacency property :

**Proposition 2.12.** *Let $V^F$ and $V^G$ be two disctinct simplices of $T_m(V)$ sharing a face of dimension $k-1$. Then there exists four points $A, D \in V^F \cap V^G$, $B \in V^F \setminus V^G$ and $C \in V^G \setminus V^F$ such that $ABCD$ forms a parallelogram, i.e.:*

$$A + D = B + C.$$

*Proof.* As $V^F$ and $V^G$ share a face of dimension $k-1$, their vertices are equal, except for one of them. Note that, in view of Remark 2.8, the common vertices appear in the same order in $V^F$ and $V^G$. Moreover, between two consecutive vertices of $V^F$ (resp. $V^G$), the Hamming distance is one. There are thus three distinct cases:

- First case: $V^F = \left[\ \mathbf{v}_0^F, \ldots, \mathbf{v}_p^F, \ldots, \mathbf{v}_k^F\right]$ and $V^G = \left[\ \mathbf{v}_0^F, \ldots, \mathbf{v}_p^G, \ldots, \mathbf{v}_k^F\right]$ with $\mathbf{v}_p^F \neq \mathbf{v}_p^G$ and $1 \leq p \leq k-1$.
  Denote by $(\mathbf{e}_0, \ldots, \mathbf{e}_k)$ the canonical basis of $\mathbb{R}^{k+1}$. Then
  $$\begin{aligned} \mathbf{v}_p^F &= \mathbf{v}_{p-1}^F + \mathbf{e}_{\sigma_F(p)} \\ \mathbf{v}_p^G &= \mathbf{v}_{p-1}^F + \mathbf{e}_{\sigma_G(p)} \\ \mathbf{v}_{p+1}^F &= \mathbf{v}_p^F + \mathbf{e}_{\sigma_F(p+1)} \\ \mathbf{v}_{p+1}^F &= \mathbf{v}_p^G + \mathbf{e}_{\sigma_G(p+1)} \\ \mathbf{v}_{\ell+1}^F &= \mathbf{v}_\ell^F + \mathbf{e}_{\sigma_F(\ell+1)} = \mathbf{v}_\ell^F + \mathbf{e}_{\sigma_G(\ell+1)} \quad \text{if } \ell \neq p, p+1. \end{aligned}$$
  Since $\mathbf{v}_p^F \neq \mathbf{v}_p^G$, we get
  $$\begin{aligned} \sigma_F(p) &= \sigma_G(p+1) \\ \sigma_F(p+1) &= \sigma_G(p) \\ \sigma_F(\ell) &= \sigma_G(\ell) \quad \text{if } \ell \neq p, p+1. \end{aligned}$$
  Setting $A = \mathbf{v}_{p-1}^F$, $B = \mathbf{v}_p^F$, $C = \mathbf{v}_p^G$ and $D = \mathbf{v}_{p+1}^F$ provides the required parallelogram, since $\mathbf{v}_{p-1}^F + \mathbf{v}_{p+1}^F = \left(\mathbf{v}_p^F - \mathbf{e}_{\sigma_F(p)}\right) + \left(\mathbf{v}_p^G + \mathbf{e}_{\sigma_G(p+1)}\right) = \mathbf{v}_p^F + \mathbf{v}_p^G$.
- Second case: $V^F = \left[\ \mathbf{v}_0^F, \mathbf{v}_1^F, \ldots, \mathbf{v}_k^F\right]$ and $V^G = \left[\ \mathbf{v}_1^F, \ldots, \mathbf{v}_k^F, \mathbf{v}_k^G\right]$. A similar argument as above shows that
  $$\begin{aligned} \sigma_F(1) &= \sigma_G(k) \\ \sigma_F(\ell+1) &= \sigma_G(\ell) \quad \text{if } 1 \leq \ell \leq k-1. \end{aligned}$$
  Setting $A = \mathbf{v}_k^F$, $B = \mathbf{v}_0^F$, $C = \mathbf{v}_k^G$ and $D = \mathbf{v}_1^F$ provides the required parallelogram.
- Third case: $V^F = \left[\ \mathbf{v}_0^F, \ldots, \mathbf{v}_{k-1}^F, \mathbf{v}_k^F\right]$ and $V^G = \left[\ \mathbf{v}_0^G, \mathbf{v}_0^F, \ldots, \mathbf{v}_{k-1}^F\right]$. A similar argument as above shows that
  $$\begin{aligned} \sigma_F(k) &= \sigma_G(1) \\ \sigma_F(\ell) &= \sigma_G(\ell+1) \quad \text{if } 1 \leq \ell \leq k-1. \end{aligned}$$
  Setting $A = \mathbf{v}_{k-1}^F$, $B = \mathbf{v}_k^F$, $C = \mathbf{v}_0^G$ and $D = \mathbf{v}_0^F$ provides the required parallelogram. □

**Corollary 2.13.** *(i) Let $V^F$ and $V^G$ be two disctinct simplices of $T_m(V)$ sharing a face of dimension $k-1$. Then there exist $|\gamma| = m-2$ and $0 \leq i < j \leq k$ such that*

$$\begin{aligned} &\mathbf{v}_{\gamma+\mathbf{e}_{i-1}+\mathbf{e}_{j-1}}(m, V) \textit{ is a vertex of } V^F \cap V^G, \\ &\mathbf{v}_{\gamma+\mathbf{e}_i+\mathbf{e}_{j-1}}(m, V) \textit{ is a vertex of } V^F \setminus V^G, \\ &\mathbf{v}_{\gamma+\mathbf{e}_{i-1}+\mathbf{e}_j}(m, V) \textit{ is a vertex of } V^G \setminus V^F, \\ &\mathbf{v}_{\gamma+\mathbf{e}_i+\mathbf{e}_j}(m, V) \textit{ is a vertex of } V^F \cap V^G. \end{aligned}$$

(ii) *Conversely, for all $|\gamma| = m-2$ and $0 \le i < j \le k$, there exist two applications $F, G : \{1,\dots,k\} \to \{1,\dots,m\}$ defining two disctinct simplices $V^F$ and $V^G$ of $T_m(V)$ sharing a face of dimension $k-1$ such that:*

$$\begin{aligned}
&\mathbf{v}_{\gamma+\mathbf{e}_{i-1}+\mathbf{e}_{j-1}}(m,V) \text{ is a vertex of } V^F \cap V^G,\\
&\mathbf{v}_{\gamma+\mathbf{e}_{i}+\mathbf{e}_{j-1}}(m,V) \text{ is a vertex of } V^F \setminus V^G,\\
&\mathbf{v}_{\gamma+\mathbf{e}_{i-1}+\mathbf{e}_{j}}(m,V) \text{ is a vertex of } V^G \setminus V^F,\\
&\mathbf{v}_{\gamma+\mathbf{e}_{i}+\mathbf{e}_{j}}(m,V) \text{ is a vertex of } V^F \cap V^G.
\end{aligned}$$

*Proof.* (i) The first part of the corollary is an easy consequence of the previous Proposition. For example, in the first case of the proof, let $i = \sigma_F(p)$ and $j = \sigma_G(p)$ and assume that $i < j$ (otherwise, exchange the role of $F$ and $G$). Write $\mathbf{v}^F_{p-1} = \mathbf{v}_\alpha(m,V)$, with $|\alpha| = m$.
Since $\mathbf{v}^F_p = \mathbf{v}^F_{p-1} + \mathbf{e}_i = \mathbf{v}_{\alpha+\mathbf{e}_i-\mathbf{e}_{i-1}}(m,V)$ and $\mathbf{v}^G_p = \mathbf{v}^F_{p-1} + \mathbf{e}_j = \mathbf{v}_{\alpha+\mathbf{e}_j-\mathbf{e}_{j-1}}(m,V)$, we can write $\alpha = \gamma + \mathbf{e}_{i-1} + \mathbf{e}_{j-1}$, with $|\gamma| = m-2$. We then get:

$$\begin{aligned}
&\mathbf{v}^F_{p-1} = \mathbf{v}^G_{p-1} = \mathbf{v}_{\gamma+\mathbf{e}_{i-1}+\mathbf{e}_{j-1}}(m,V) \text{ is a vertex of } V^F \cap V^G,\\
&\mathbf{v}^F_{p} = \mathbf{v}_{\gamma+\mathbf{e}_{i}+\mathbf{e}_{j-1}}(m,V) \text{ is a vertex of } V^F \setminus V^G,\\
&\mathbf{v}^G_{p} = \mathbf{v}_{\gamma+\mathbf{e}_{i-1}+\mathbf{e}_{j}}(m,V) \text{ is a vertex of } V^G \setminus V^F,\\
&\mathbf{v}^F_{p+1} = \mathbf{v}^G_{p+1} = \mathbf{v}_{\gamma+\mathbf{e}_{i}+\mathbf{e}_{j}}(m,V) \text{ is a vertex of } V^F \cap V^G,
\end{aligned}$$

as wanted.
The other cases are similar.

(ii) We now prove the second assertion.

Assume first that $i \ge 1$.

Let $f_p$ $(1 \le p \le k)$ be the numbers such that

$$\gamma + \mathbf{e}_0 + \mathbf{e}_{j-1} = (f_1, f_2 - f_1, \dots, f_k - f_{k-1}, d - f_k).$$

Then $1 \le f_1 \le \cdots \le f_k \le m$.
Moreover, define the permutations $\sigma_F$ and $\sigma_G$ as follows:

$$\begin{cases}
\sigma_F(\ell) = \sigma_G(\ell) = \ell \text{ if } 1 \le \ell \le i-1\\
\sigma_F(i) = i \text{ and } \sigma_G(i) = j\\
\sigma_F(i+1) = j \text{ and } \sigma_G(i+1) = i\\
\sigma_F(\ell+1) > \sigma_F(\ell) \text{ and } \sigma_G(\ell+1) > \sigma_G(\ell) \text{ if } i+2 \le \ell \le k-1.
\end{cases}$$

Defining the application $F : \{1,\dots,k\} \to \{1,\dots,m\}$ by $F(j) = f_{\sigma_F(j)}$, we can construct the simplex $V^F = \left[\mathbf{v}^F_0, \dots, \mathbf{v}^F_k\right]$ of $T_m(V)$, whose vertices satisfy:

$$\begin{cases} \mathbf{v}_0^F = \mathbf{v}_{\gamma+\mathbf{e}_0+\mathbf{e}_{j-1}}(m,V) \\ \mathbf{v}_{i-1}^F = \mathbf{v}_{\gamma+\mathbf{e}_{i-1}+\mathbf{e}_{j-1}}(m,V) \\ \mathbf{v}_i^F = \mathbf{v}_{\gamma+\mathbf{e}_i+\mathbf{e}_{j-1}}(m,V) \\ \mathbf{v}_{i+1}^F = \mathbf{v}_{\gamma+\mathbf{e}_i+\mathbf{e}_j}(m,V). \end{cases}$$

Similarly, if $G(j) = f_{\sigma_G(j)}$, we get the simplex $V^G$ of $T_m(V)$, with:

$$\begin{cases} \mathbf{v}_0^G = \mathbf{v}_{\gamma+\mathbf{e}_0+\mathbf{e}_{j-1}}(m,V) \\ \mathbf{v}_{i-1}^G = \mathbf{v}_{\gamma+\mathbf{e}_{i-1}+\mathbf{e}_{j-1}}(m,V) \\ \mathbf{v}_i^G = \mathbf{v}_{\gamma+\mathbf{e}_{i-1}+\mathbf{e}_j}(m,V) \\ \mathbf{v}_{i+1}^G = \mathbf{v}_{\gamma+\mathbf{e}_i+\mathbf{e}_j}(m,V), \end{cases}$$

as wanted.

The case $i = 0$ is similar. □

**Subtriangulations** We next consider triangulating the simplices in $T_m(V)$, which will be needed in section 3.

**Proposition 2.14.** *Let $m, n \geq 2$, and consider a simplex $V$. Then the standard triangulation $T_{mn}(V)$ is exactly the collection $\left(T_n(V^F)\right)_{V^F \in T_m(V)}$.*

*Proof.* ◦ We can assume, without loss of generality, that $V = \Lambda = mV^{\mathrm{Id}}$.

◦ Consider a simplex $V^F = \mathbf{v}_0^F + V^{\sigma_F} \in T_m(\Lambda)$. Let $W \in T_n(V^F)$.
Consider the linear application $u : \mathbb{R}^k \to \mathbb{R}^k$ permuting (and scaling) the canonical basis $(\mathbf{e}_1, \ldots, \mathbf{e}_k)$ of $\mathbb{R}^k$ as follows:

$$u(\mathbf{e}_i) = \frac{\mathbf{e}_{\sigma_F(i)}}{n} \quad (1 \leq i \leq k).$$

Then $W \in T_n(V^F) = T_n\left(\mathbf{v}_0^F + u\left(nV^{\mathrm{Id}}\right)\right)$. Then $W$ can be written in the following form :

$$W = \mathbf{v}_0^F + u\left(\mathbf{v}_0^G + V^{\sigma_G}\right) = \mathbf{v}_0^F + u\left(\mathbf{v}_0^G\right) + u\left(V^{\sigma_G}\right),$$

where $\mathbf{v}_0^G + V^{\sigma_G}$ is a simplex of the standard triangulation $T_n\left(nV^{\mathrm{Id}}\right)$.
Obviously, $u\left(V^{\sigma_G}\right) = \frac{1}{n}V^{\sigma_F \circ \sigma_G}$. Moreover, an easy computation shows that

$$\mathbf{v}_0^F + u\left(\mathbf{v}_0^G\right) = \frac{1}{n}(mn - h_1, \ldots, mn - h_k),$$

where $h_\ell = nf_\ell + g_{\sigma_F^{-1}(\ell)} - g_{\sigma_F^{-1}(\ell)+1} \in \{1, \ldots, mn\}$. We can now define an application $H : \{1, \ldots, k\} \to \{1, \ldots, mn\}$ by:

$$H(i) = h_{\sigma_F \circ \sigma_G(i)} \quad (1 \leq i \leq k).$$

Then $W = \frac{1}{n}(\mathbf{v}_0^H + V^{\sigma_H}) \in \frac{1}{n} T_{mn}(mnV^{\mathrm{Id}}) = T_{mn}(mV^{\mathrm{Id}})$.
Hence, every simplex of the form $T_n(V^F)$ with $V^F \in T_m(\Lambda)$ belongs to the standard triangulation $T_{mn}(\Lambda)$.

- It remains to note that both collections have exactly $(mn)^k$ simplices to conclude. □

# 3 Approximation by the Bernstein coefficients

## 3.1 Notations

In this section, we state a quantitative result concerning the coefficients of a real polynomial, expressed in the Bernstein basis with respect to a simplex. We provide an explicit bound on the gap between these coefficients and the graph of the considered polynomial. This generalizes known results in dimensions 1 ([NPL]) and 2 ([R]).

Let $f \in \mathbb{R}[\mathbf{X}] = \mathbb{R}[X_1, \ldots, X_k]$ be a polynomial of degree $\leq d$, and $V = [\mathbf{v}_0, \ldots, \mathbf{v}_k]$ a simplex of $\mathbb{R}^k$. Let

$$f = \sum_{|\alpha|=d} b_\alpha(f, d, V) B_\alpha^d$$

be the expansion of $f$ in the Bernstein basis of degree $d$ w.r.t. $V$.

The Bernstein coefficients of a polynomial $f$ already some geometric information, which can be expressed in terms of the so-called control points of $f$:

**Definition 3.1.** *Let $V = [\mathbf{v}_0, \ldots, \mathbf{v}_k]$ be a non-degenerate simplex of $\mathbb{R}^k$, $f \in \mathbb{R}[\mathbf{X}]$ a polynomial of degree $\leq d$ and $b_\alpha(f, d, V)$ ($|\alpha| = d$) its Bernstein coefficients of degree $d$ w.r.t. $V$.*

- *The grid points of degree $d$ associated to $V$ are the points*

$$\mathbf{v}_\alpha(d, V) = \frac{\alpha_0 \mathbf{v}_0 + \cdots + \alpha_k \mathbf{v}_k}{d} \in \mathbb{R}^k \quad (|\alpha| = d).$$

- *The control points associated to $f$ of degree $d$ w.r.t. $V$ are the points*

$$\mathbf{C}_\alpha = \big(\mathbf{v}_\alpha(d, V), b_\alpha(f, d, V)\big) \in \mathbb{R}^{k+1} \quad (|\alpha| = d).$$

  *The control points of $f$ form its control net of degree $d$.*

- *The discrete graph of $f$ of degree $d$ w.r.t. $V$ is formed by the points $\Big(\mathbf{v}_\alpha(d, V), f\big(\mathbf{v}_\alpha(d, V)\big)\Big)_{|\alpha|=d}$.*

We then have the following properties:

**Proposition 3.2.** *Keeping the same notations, we have:*

*(i)* **linear precision***: if* $\deg f \leq 1$*, then:*

$$\forall |\alpha| = d,\ b_\alpha(f, d, V) = f(\mathbf{v}_\alpha(d, V)).$$

*(ii)* **interpolation at the vertices***: if* $(\mathbf{e}_0, \dots, \mathbf{e}_k)$ *denotes the canonical basis of* $\mathbb{R}^k$*, then:*

$$\forall i \in \{0, \dots, k\},\ b_{d\mathbf{e}_i} = f(\mathbf{v}_i).$$

*(iii)* **convex hull property***: the graph of* $f$ *is contained in the convex hull of its associated control points.*

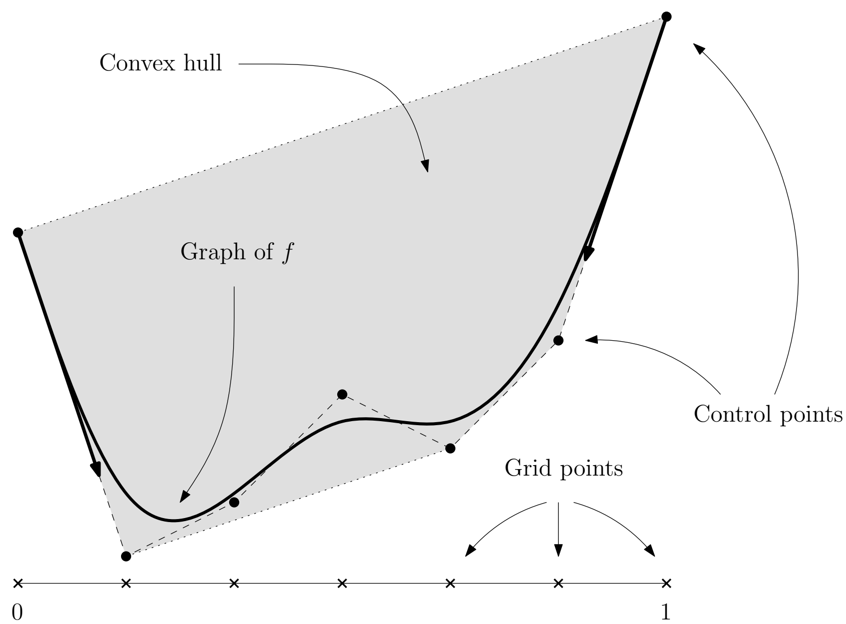



Figure 4: Example of a control net of degree 6 in dimension 1.

We aim at comparing the discrete graph of $f$ to its control net.

**Remark 3.3.** *If* $\deg f \leq 1$*, the control net is exactly the discrete graph of* $f$*, according to the linear precision property.*

## 3.2 Control polytope

In order to compare the control net to the discrete graph of a polynomial, we first introduce the (continuous) notion of control polytope:

**Definition 3.4.** *The control polytope associated to $f$ of degree $d$ w.r.t. a simplex $V = [\mathbf{v}_0, \ldots, \mathbf{v}_k]$ of $\mathbb{R}^k$ is the unique continuous function $\hat{f}$, linear on each simplex of the standard triangulation $T_d(V)$ satisfying the following interpolation property:*

$$\forall |\alpha| = d, \ \hat{f}(\mathbf{v}_\alpha(d, V)) = b_\alpha(f, d, V).$$

We will mainly study the convexity of the control polytope. In order to do so, we introduce the notion of second differences:

**Definition 3.5.** *Let $(\mathbf{e}_0, \ldots, \mathbf{e}_k)$ be the standard basis of $\mathbb{R}^{k+1}$ (with the convention $\mathbf{e}_{-1} = \mathbf{e}_k$), and $V = [\mathbf{v}_0, \ldots, \mathbf{v}_k]$ a simplex of $\mathbb{R}^k$.*
*For $|\gamma| = d - 2$ et $0 \leq i < j \leq k$, define the quantity (where $b_\alpha$ stands for $b_\alpha(f, d, V)$):*

$$\nabla^2 b_{\gamma,i,j}(f, d, V) = b_{\gamma+\mathbf{e}_i+\mathbf{e}_{j-1}} + b_{\gamma+\mathbf{e}_{i-1}+\mathbf{e}_j} - b_{\gamma+\mathbf{e}_{i-1}+\mathbf{e}_{j-1}} - b_{\gamma+\mathbf{e}_i+\mathbf{e}_j}.$$

*The collection $\nabla^2 b(f, d, V) = \left(\nabla^2 b_{\gamma,i,j}(f, d, V)\right)_{\substack{|\gamma|=d-2 \\ 0\leq i<j\leq k}}$ forms the second differences of $f$ of degree $d$ w.r.t. $V$.*

**Notation 3.6.** *Let $\|\nabla_2 b(f, d, V)\|_\infty$ denote the quantity*

$$\max_{\substack{|\gamma|=d-2 \\ 0\leq i<j\leq k}} |\nabla^2 b_{\gamma,i,j}(f, d, V)|.$$

**Example 3.7.** *Let $q(X_1, \ldots, X_k)$ be a quadratic form, whose associated matrix is denoted by $M = (m_{ij})_{1\leq i,j\leq k}$. Let $d \geq 2$ be an integer, and $\Delta$ be the standard simplex of $\mathbb{R}^k$:*

$$\Delta = \{(x_1, \ldots, x_k) \in \mathbb{R}^k_+ \mid \sum x_i \leq 1\}.$$

*Then:*

$$\forall \ |\alpha| = d, \ b_\alpha(q, d, \Delta) = \frac{1}{d(d-1)} \left( q(\alpha_1, \ldots, \alpha_k) - \sum_{i=1}^{k} m_{ii}\alpha_i \right).$$

*The second differences of $q$ are easily computed:*

$$\nabla^2 b_{\gamma,i,j}(q, d, \Delta) = \frac{2}{d(d-1)} \left(m_{i-1,j} + m_{i,j-1} - m_{i,j} - m_{i-1,j-1}\right), \tag{3.1}$$

*with the conventions $m_{0,j} = m_{i,0} = 0$, $m_{-1,j} = m_{k,j}$ et $m_{-1,-1} = m_{k,k}$.*

The following lemma is fundamental:

**Lemma 3.8.** *There exists a unique quadratic form $q_d(X_1, \ldots, X_k)$ such that*

$$\nabla^2 b(q_d, d, \Delta) = (1, \ldots, 1).$$

*Moreover, its associated matrix $N$ is given by $N = \dfrac{d(d-1)}{2} M$, where $M = (m_{ij})_{1\leq i,j\leq k}$ is the symmetric matrix defined by:*

$$m_{i,j} = i(k - j + 1) \ \text{if} \ i \leq j. \tag{3.2}$$

*Proof.* The existence is a straightforward computation, while the relations (3.1) easily implies the unicity of such a quadratic form. □

**Remark 3.9.** *The matrix $M$ in the previous lemma is positive definite. Indeed, if $P$ denotes the matrix*

$$P = \begin{pmatrix} 1 & -1 & & \\ & \ddots & \ddots & \\ & & \ddots & -1 \\ & & & 1 \end{pmatrix},$$

*then*

$$P^t M P = (k+1)I - J$$

*where $P^t$ denotes the transpose of the matrix $P$, $I$ is the identity matrix, and $J$ is the matrix whose coefficients are all equal to 1. It is easy to show that $(k+1)I - J$ is positive definite, and so is $M$.*

The second differences can be interpreted in terms of the so-called directional derivatives.

Indeed, defining the directional derivative $D_{\mathbf{u}}$ ($\mathbf{u} \in \mathbb{R}^k$) by $D_{\mathbf{u}} = \sum_{i=1}^{k} u_i \dfrac{\partial}{\partial x_i}$, we have, for all $0 \le i < j \le k$ and $|\gamma| = d-2$:

$$b_\gamma\left(D_{\mathbf{v}_j - \mathbf{v}_{j-1}} D_{\mathbf{v}_i - \mathbf{v}_{i-1}} f, d-2, V\right) = -d(d-1)\nabla^2 b_{\gamma,i,j}(f,d,V). \tag{3.3}$$

**Remark 3.10.** *Assume that $(\mathbf{u}_1, \ldots, \mathbf{u}_k)$ form a basis of $\mathbb{R}^k$.*
*If $D_{\mathbf{u}_i} f$ is the zero polynomial for all $i$, then the gradient of $f$ is orthogonal to each $\mathbf{u}_i$. Hence, the gradient of $f$ is zero, and $f$ is constant.*
*Similarly, if $D_{\mathbf{u}_i} f$ is constant for all $i$, then the gradient of $f$ is constant, and the degree of $f$ is $\le 1$.*

We can now express the convexity of $\hat{f}$ in terms of the second differences:

**Theorem 3.11.** *The control polytope $\hat{f}$ is convex on $V$ if and only if $\nabla^2 b(f,d,V) \le 0$ (where the inequality is meant componentwise).*

*Proof.* We first note that $\hat{f}$ is convex on $V$ if and only if for any line $L$, the restriction $\hat{f}|_{L\cap V}$ is convex. By a continuity argument, it is sufficient to consider the lines $L$ such that :

- for all subsimplex $U \in T_d(V)$, $L \cap U$ is not contained in a face of $U$ of dimension $< k$
- no grid point $\mathbf{v}_\alpha(d,V)$ is contained in $L$.

But this is equivalent to the convexity of $\hat{f}$ restricted to the union of each pair of subsimplices in $T_d(V)$ that share a common face of dimension $k-1$.

• Assume that $\hat{f}$ is convex, and consider some $|\gamma| = d-2$ and $0 \leq i < j \leq k$. According to Corollary 2.13, there exist two simplices $V^F$ and $V^G$ sharing a face of dimension $k-1$ such that

$$\begin{aligned}
&\mathbf{v}_{\gamma+\mathbf{e}_{i-1}+\mathbf{e}_{j-1}}(m,V) \text{ is a vertex of } V^F \cap V^G,\\
&\mathbf{v}_{\gamma+\mathbf{e}_{i}+\mathbf{e}_{j-1}}(m,V) \text{ is a vertex of } V^F \setminus V^G,\\
&\mathbf{v}_{\gamma+\mathbf{e}_{i-1}+\mathbf{e}_{j}}(m,V) \text{ is a vertex of } V^G \setminus V^F,\\
&\mathbf{v}_{\gamma+\mathbf{e}_{i}+\mathbf{e}_{j}}(m,V) \text{ is a vertex of } V^F \cap V^G.
\end{aligned}$$

We then have (with some slight abuse of notation):

$$\begin{aligned}
b_{\gamma+\mathbf{e}_{i-1}+\mathbf{e}_j} + b_{\gamma+\mathbf{e}_i+\mathbf{e}_{j-1}} &= \hat{f}(\mathbf{v}_{\gamma+\mathbf{e}_{i-1}+\mathbf{e}_j}) + \hat{f}(\mathbf{v}_{\gamma+\mathbf{e}_i+\mathbf{e}_{j-1}})\\
&= 2\hat{f}\left(\frac{\mathbf{v}_{\gamma+\mathbf{e}_{i-1}+\mathbf{e}_j} + \mathbf{v}_{\gamma+\mathbf{e}_i+\mathbf{e}_{j-1}}}{2}\right)\\
&\qquad\qquad\qquad \text{since } \hat{f} \text{ is linear on } V^F \cap V^G\\
&= 2\hat{f}\left(\frac{\mathbf{v}_{\gamma+\mathbf{e}_{i-1}+\mathbf{e}_{j-1}} + \mathbf{v}_{\gamma+\mathbf{e}_i+\mathbf{e}_j}}{2}\right)\\
&\leq \hat{f}(\mathbf{v}_{\gamma+\mathbf{e}_{i-1}+\mathbf{e}_{j-1}}) + \hat{f}(\mathbf{v}_{\gamma+\mathbf{e}_i+\mathbf{e}_j}) \text{ since } \hat{f} \text{ is convex}\\
&= b_{\gamma+\mathbf{e}_{i-1}+\mathbf{e}_{j-1}} + b_{\gamma+\mathbf{e}_i+\mathbf{e}_j}.
\end{aligned}$$

In other words,

$$\nabla^2 b_{\gamma,i,j}(f,d,V) = b_{\gamma+\mathbf{e}_i+\mathbf{e}_{j-1}} + b_{\gamma+\mathbf{e}_{i-1}+\mathbf{e}_j} - b_{\gamma+\mathbf{e}_{i-1}+\mathbf{e}_{j-1}} - b_{\gamma+\mathbf{e}_i+\mathbf{e}_j} \leq 0.$$

We have just shown that if $\hat{f}$ is convex, then $\nabla^2 b(f,d,V) \leq 0$.

• Conversely, assume that $\nabla^2 b(f,d,V) \leq 0$. We now prove that this implies the convexity of $\hat{f}$. It is sufficient to prove the convexity of $\hat{f}$ restricted to the union of each pair of subsimplices in $T_d(V)$ that share a common face of dimension $k-1$. Let $V^F$ and $V^G$ be such a pair of simplices. According to Corollary 2.13, there exist $|\gamma| = d-2$ and $0 \leq i < j \leq k$ such that

$$\begin{aligned}
&\mathbf{v}_{\gamma+\mathbf{e}_{i-1}+\mathbf{e}_{j-1}}(m,V) \text{ is a vertex of } V^F \cap V^G,\\
&\mathbf{v}_{\gamma+\mathbf{e}_{i}+\mathbf{e}_{j-1}}(m,V) \text{ is a vertex of } V^F \setminus V^G,\\
&\mathbf{v}_{\gamma+\mathbf{e}_{i-1}+\mathbf{e}_{j}}(m,V) \text{ is a vertex of } V^G \setminus V^F,\\
&\mathbf{v}_{\gamma+\mathbf{e}_{i}+\mathbf{e}_{j}}(m,V) \text{ is a vertex of } V^F \cap V^G.
\end{aligned}$$

By a standard convexity argument, it is sufficient to prove that $\hat{f}$ is "midpoint convex" on $V^F \cup V^G$, *i.e.*:

$$\forall x, y \in V^F \cup V^G,\ \hat{f}\left(\frac{x+y}{2}\right) \leq \frac{\hat{f}(x) + \hat{f}(y)}{2}.$$

It is easy to check that this is equivalent to (again with a slight abuse of notation):

$$\hat{f}\left(\frac{\mathbf{v}_{\gamma+\mathbf{e}_{i-1}+\mathbf{e}_{j-1}} + \mathbf{v}_{\gamma+\mathbf{e}_i+\mathbf{e}_j}}{2}\right) \leq \frac{\hat{f}(\mathbf{v}_{\gamma+\mathbf{e}_i+\mathbf{e}_{j-1}}) + \hat{f}(\mathbf{v}_{\gamma+\mathbf{e}_{i-1}+\mathbf{e}_j})}{2},$$

or equivalently, since $\hat{f}$ is linear on $V^F \cap V^G$:

$$\hat{f}(\mathbf{v}_{\gamma+\mathbf{e}_{i-1}+\mathbf{e}_{j-1}}) + \hat{f}(\mathbf{v}_{\gamma+\mathbf{e}_i+\mathbf{e}_j}) - \hat{f}(\mathbf{v}_{\gamma+\mathbf{e}_i+\mathbf{e}_{j-1}}) - \hat{f}(\mathbf{v}_{\gamma+\mathbf{e}_{i-1}+\mathbf{e}_j}) \leq 0,$$

that it to say:

$$\nabla^2 b_{\gamma,i,j}(f, d, V) \leq 0,$$

which is true by hypothesis. □

## 3.3 Main result

We now state the main result of the section, bounding the gap between the control net and the discrete graph of $f$:

**Theorem 3.12.** *With the previous notations, we have:*

$$\max_{|\alpha|=d} \left| f\big(\mathbf{v}_\alpha(d, V)\big) - b_\alpha(f, d, V)\right| \leq C(k, d) \left\|\nabla^2 b(f, d, V)\right\|_\infty,$$

*where*

$$C(k, d) = \frac{\left\lfloor \frac{d^2 k(k+2)}{12} \right\rfloor}{2d}.$$

*Proof.* ∘ Without loss of generality, we can assume that $V$ is the standard simplex $\Delta = [\mathbf{0}, \mathbf{e}_1, \ldots, \mathbf{e}_k]$, where $(\mathbf{e}_1, \ldots, \mathbf{e}_k)$ denotes the standard basis of $\mathbb{R}^k$.

Indeed, let $h$ denote the affine transformation mapping $\mathbf{0}$ on $\mathbf{v}_0$ and each $\mathbf{e}_i$ on $\mathbf{v}_i$. Define the polynomial $g = f \circ h$. Since the barycentric coordinates are preserved under affine transformation, we have

$$b(g, d, \Delta) = b(f, d, V).$$

Moreover, Equation (3.3) shows that

$$\nabla^2 b(g, d, \Delta) = \nabla^2 b(f, d, V).$$

Hence, we can assume that $V = \Delta$.

∘ We now claim that the worst case is achieved by the unique quadratic form $q_d$ defined in Example 3.8 and satisfying

$$\nabla^2 b(q_d, d, \Delta) = (1, \ldots, 1).$$

More precisely, we show the following inequality:

$$\forall|\alpha| = d,\ \left|\left(f - \hat{f}\right)(\mathbf{v}_\alpha(d,V))\right| \le (q_d - \hat{q_d})(\mathbf{v}_\alpha(d,V))\left\|\nabla^2 b(f,d,\Delta)\right\|_\infty .$$

If $\left\|\nabla^2 b(f,d,\Delta)\right\|_\infty = 0$, the result is clear, since in this case the degree of $f$ is $\le 1$.

Else we can assume that $f$ is normalized, such that $\left\|\nabla^2 b(f,d,\Delta)\right\|_\infty = 1$. Since $\nabla^2 b(q_d,d,\Delta) = (1,\dots,1)$, we have $\nabla^2 b(q_d + f,d,\Delta) \ge 0$. The control polytope of $q_d + f$ is thus convex according to Theorem 3.11. Hence:

$$\begin{aligned}(q_d + f)(\mathbf{v}_\alpha(d,\Delta)) &= \sum_{|\beta|=d} b_\beta(q_d + f,d,\Delta)B^d_\beta(\mathbf{v}_\alpha(d,\Delta))\\ &= \sum_{|\beta|=d} (\widehat{q_d + f})(\mathbf{v}_\beta(d,\Delta))B^d_\beta(\mathbf{v}_\alpha(d,\Delta))\\ &\ge (\widehat{q_d + f})\left(\sum_{|\beta|=d} B^d_\beta(\mathbf{v}_\alpha(d,\Delta))\mathbf{v}_\beta(d,\Delta)\right),\end{aligned}$$

where the last inequality comes from the convexity of $(\widehat{q_d + f})$.
The linear precision property recalled in Remark 3.3 implies that

$$\sum_{|\beta|=d} B^d_\beta(\mathbf{v}_\alpha(d,\Delta))\mathbf{v}_\beta(d,\Delta) = \mathbf{v}_\alpha(d,\Delta).$$

Hence, $(q_d + f)(\mathbf{v}_\alpha(d,\Delta)) \ge (\widehat{q_d + f})(\mathbf{v}_\alpha(d,\Delta))$, that is to say:

$$(q_d - \hat{q_d})(\mathbf{v}_\alpha(d,\Delta)) \ge -\left(f - \hat{f}\right)(\mathbf{v}_\alpha(d,\Delta)).$$

Analogously, considering $q_d - f$, we get:

$$(q_d - \hat{q_d})(\mathbf{v}_\alpha(d,\Delta)) \ge \left(f - \hat{f}\right)(\mathbf{v}_\alpha(d,\Delta)).$$

Hence,

$$\left|\left(f - \hat{f}\right)(\mathbf{v}_\alpha(d,\Delta))\right| \le (q_d - \hat{q_d})(\mathbf{v}_\alpha(d,\Delta)),$$

as announced.

- It remains to show that

$$\max_{|\alpha|=d}(q_d - \hat{q_d})(\mathbf{v}_\alpha(d,\Delta)) \le \frac{\left\lfloor \frac{d^2k(k+2)}{12}\right\rfloor}{2d}.$$

According to Example 3.7, we have:

$$\begin{aligned}(q_d - \hat{q_d})(\mathbf{v}_\alpha(d,\Delta)) &= q_d(\mathbf{v}_\alpha(d,\Delta)) - b_\alpha(q_d,d,\Delta)\\ &= \Big(-\frac{1}{2d}\Big)(\alpha_1,\dots,\alpha_k)M(\alpha_1,\dots,\alpha_k)^t + \frac{1}{2}\sum_{i=1}^{k} m_{ii}\alpha_i.\end{aligned}$$

Define the application gap : $\mathbb{R}^k \to \mathbb{R}$ by:

$$\mathrm{gap}(x_1,\dots,x_k) = \Big(-\frac{1}{2d}\Big)(x_1,\dots,x_k)M(x_1,\dots,x_k)^t + \frac{1}{2}\sum_{i=1}^{k} m_{ii}x_i.$$

gap is a $C^\infty$ function, whose hessian matrix is equal to $-M/d$ and is thus negative definite according to Remark 3.9. Hence, gap is strictly concave. Its gradient is zero at the point $\left(\frac{d}{k+1},\dots,\frac{d}{k+1}\right)$, and

$$\mathrm{gap}\left(\frac{d}{k+1},\dots,\frac{d}{k+1}\right) = \frac{dk(k+2)}{24}$$

is then its global maximum.
As $2d\,\mathrm{gap}(\alpha_1,\dots,\alpha_k)$ is an integer for all $\alpha \in \mathbb{N}^{k+1}$, we get:

$$\forall |\alpha| = d,\ 2d\,\mathrm{gap}(\alpha_1,\dots,\alpha_k) \leq \left\lfloor \frac{d^2k(k+2)}{12} \right\rfloor.$$

Combining the previous results leads to the desired conclusion. $\square$

**Remark 3.13.** *The bound in theorem 3.12 is sharp, since it is achieved by the quadratic form $q_d$.*

**Remark 3.14.** *This result generalizes the result in [NPL] (univariate case), as well as the result in [R] (bivariate case).*

**Remark 3.15.** *In contrast with its proof, which uses the control polytope of $f$, the statement of Theorem 3.12 involves only its Bernstein coefficients, independently of the choice of a triangulation of $V$.*
*A continuous version of this result can be found in [L].*

# 4 Convergence of the control net to the discrete graph

In this section, we show that the control points of $f$ converge to its graph at the grid points. Two kinds of convergence are considered : by degree elevation and by subdivision. Without loss of generality, we can assume that the polynomial $f$ is studied on the standard simplex $\Delta$.

## 4.1 Convergence by degree elevation

Every polynomial $f \in \mathbb{R}[\mathbf{X}] = \mathbb{R}[X_1,\dots,X_k]$ of degree $\leq d$ can be expressed in the Bernstein basis of degree $D$, for each $D > d$: this is called degree elevation.

We first study the behaviour of the second differences under degree elevation :

**Lemma 4.1.** *Let $f \in \mathbb{R}[\mathbf{X}]$ be a polynomial of degree $d$ on the standard simplex $\Delta$. Let $D > d$. Then :*

$$\left\|\nabla^2 b(f, D, \Delta)\right\|_\infty \leq \frac{d(d-1)}{D(D-1)} \left\|\nabla^2 b(f, d, \Delta)\right\|_\infty .$$

*Proof.*

- The result is clear if $\left\|\nabla^2 b(f, d, \Delta)\right\|_\infty = 0$, since in this case, Equation 3.3 shows that the directional derivatives are zero, and thus their expansion in every Bernstein basis of degree $\geq D-2$ is null.
- First, assume that $\left\|\nabla^2 b(f, d, \Delta)\right\|_\infty = 1$.
  Consider the quadratic form $q_d$ given in Example 3.8, satisfying

  $$\nabla^2 b(q_d, d, \Delta) = (1, \dots, 1).$$

  Then $\nabla^2 b(q_d \pm f, d, \Delta) \geq 0$, or equivalently:

  $$b_\gamma \left(D_{\Delta_j - \Delta_{j-1}} D_{\Delta_i - \Delta_{i-1}} (q_d \pm f), d-2, \Delta\right) \leq 0,$$

  for every $|\gamma| = d-2$ and $0 \leq i < j \leq k$.
  But then, for every $|\gamma| = d-2$ and $0 \leq i < j \leq k$, we have

  $$b_\gamma \left(D_{\Delta_j - \Delta_{j-1}} D_{\Delta_i - \Delta_{i-1}} (q_d \pm f), D-2, \Delta\right) \leq 0.$$

  Indeed, the latter coefficients are computed by degree elevation, which involves only nonnegative combinations of the former coefficients.
  We then have

  $$\nabla^2 b(q_d \pm f, D, \Delta) \geq 0,$$

  from which we get

  $$\left\|\nabla^2 b(f, D, \Delta)\right\|_\infty \leq \left\|\nabla^2 b(q_d, D, \Delta)\right\|_\infty .$$

- In the general case, we get:

  $$\left\|\nabla^2 b(f, D, \Delta)\right\|_\infty \leq \left\|\nabla^2 b(q_d, D, \Delta)\right\|_\infty \left\|\nabla^2 b(f, d, \Delta)\right\|_\infty .$$

- Let $q_D$ be the quadratic form verifying

  $$\nabla^2 b(q_D, D, \Delta) = (1, \dots, 1).$$

  Then $q_d = \dfrac{d(d-1)}{D(D-1)} q_D$, so that

  $$\nabla^2 b(q_d, D, \Delta) = \frac{d(d-1)}{D(D-1)} (1, \dots, 1).$$

  Consequently,

  $$\left\|\nabla^2 b(q_d, D, \Delta)\right\|_\infty = \frac{d(d-1)}{D(D-1)},$$

and then

$$\left\|\nabla^2 b(f,D,\Delta)\right\|_\infty \leq \frac{d(d-1)}{D(D-1)} \left\|\nabla^2 b(f,d,\Delta)\right\|_\infty ,$$

as announced. □

We can now deduce the following theorem, expressing the convergence of the control points under degree elevation:

**Theorem 4.2.** *Let $f \in \mathbb{R}[\mathbf{X}]$ be a polynomial of degree $d$ over the standard simplex $\Delta$. Let $D > d$. Then:*

$$\max_{|\alpha|=D} |f(\mathbf{v}_\alpha(D,\Delta)) - b_\alpha(f,D,\Delta)| \leq \frac{k(k+2)}{24} \frac{d(d-1)}{D-1} \left\|\nabla^2 b(f,d,\Delta)\right\|_\infty .$$

*Proof.* Theorem 3.12 implies that

$$\max_{|\alpha|=D} |f(\mathbf{v}_\alpha(D,\Delta)) - b_\alpha(f,D,\Delta)| \leq \frac{Dk(k+2)}{24} \left\|\nabla^2 b(f,D,\Delta)\right\|_\infty .$$

Besides, the previous lemma implies that

$$\left\|\nabla^2 b(f,D,\Delta)\right\|_\infty \leq \frac{d(d-1)}{D(D-1)} \left\|\nabla^2 b(f,d,\Delta)\right\|_\infty .$$

We then obtain

$$\max_{|\alpha|=D} |f(\mathbf{v}_\alpha(D,\Delta)) - b_\alpha(f,D,\Delta)| \leq \frac{Dk(k+2)}{24} \frac{d(d-1)}{D(D-1)} \left\|\nabla^2 b(f,d,\Delta)\right\|_\infty ,$$

as announced. □

**Remark 4.3.** *Note that, by affine transformation, the result holds for every simplex $V \subset \mathbb{R}^k$.*

**Remark 4.4.** *Note that the previous theorem not only express the convergence of the control points under degree elevation, but also shows that the rate is linear in $1/D$. This result was already known ([KP]). The interest of Theorem 4.9 lies in the fact that the bound is explicit in terms of the dimension, the degree and the Bernstein expansion of $f$ of degree $d$ over $\Delta$.*

## 4.2 Convergence under subdivision

Let $f \in \mathbb{R}[\mathbf{X}]$ be a polynomial of degree $d$, over the standard simplex $\Delta$. Assume that $\Delta$ has been subdivided, *i.e.*

$$\Delta = U^1 \cup \cdots \cup U^s,$$

where the interiors of the simplices $U^i$ $(1 \leq i \leq s)$ are disjoint. Note that a triangulation of $\Delta$ is a particular case of subdivision.
The expansion of $f$ in the Bernstein basis of degree $d$ associated with each subsimplex $U^i$ can be computed using only convex combinations of its Bernstein coefficients w.r.t $\Delta$. This can be done by successive calls to the De Casteljau algorithm ([P]), which we recall for the readers' convenience.

**Notation 4.5.** *If $V = [\mathbf{v}_0, \ldots, \mathbf{v}_k]$ is a simplex of $\mathbb{R}^k$ and $M \in \mathbb{R}^k$, the simplices $V^{[i]}$ $(i = 0, \ldots, k)$ are defined as follows:*

$$V^{[i]} = [\mathbf{v}_0, \ldots, \mathbf{v}_{i-1}, M, \mathbf{v}_{i+1}, \ldots, \mathbf{v}_k].$$

*In what follows, if $\alpha \in \mathbb{N}^{k+1}$ and $0 \leq i \leq k$, we write*

$$\widehat{\alpha_i} = (\alpha_0, \ldots, \alpha_{i-1}, 0, \alpha_{i+1}, \ldots, \alpha_k).$$

*Recall that the barycentric coordinates of $M$ w.r.t. $V$ are denoted by*

$$(\lambda_0(M), \ldots, \lambda_k(M)).$$

*The standard basis of $\mathbb{R}^{k+1}$ is denoted by $(\mathbf{e}_0, \ldots, \mathbf{e}_k)$.*

**Algorithm 4.6** (De Casteljau)**.**

***Require:*** *a simplex $V$, the Bernstein expansion $b(f, d, V)$ of a polynomial $f$ of degree $d$ over $V$, and a point $M \in \mathbb{R}^k$.*

***Ensure:*** *the Bernstein expansions $b(f, d, V^{[i]})$ of $f$ associated to the simplices $V^{[i]}$.*

***Initialization*** *: $\forall |\alpha| = d,\ b_\alpha^{(0)} := b_\alpha(f, d, V)$.*

***for*** $l = 1, \ldots, d$ ***do***
  ***for*** $|\alpha| = d - l$ ***do***
    *Compute* $b_\alpha^{(l)} := \sum_{p=0}^{k} \lambda_p(M) b_{\alpha + \mathbf{e}_p}^{(l-1)}$
  ***end for***
***end for***

*Return* $b_\alpha(f, d, V^{[i]}) = b_{\widehat{\alpha_i}}^{(\alpha_i)}$ $(|\alpha| = d,\ i = 0, \ldots, k)$.

**Remark 4.7.** *If $U = [\mathbf{u}_0, \ldots, \mathbf{u}_k]$ is a subsimplex of $V$, then the Bernstein expansion $b(f, d, U)$ can be computed from $b(f, d, V)$ by $k + 1$ successive calls to De Casteljau's algorithm at the points $\mathbf{u}_0, \ldots, \mathbf{u}_k$. In each call, only convex combinations of the Bernstein coefficients $b(f, d, V)$ are involved. This process is called reparametrization.*

We now study the behaviour of the second differences of $f$ under subdivision of the standard simplex $\Delta$.

Consider a subsimplex $U \subset \Delta$. Denote by $h$ the diameter of $U$. We then have:

**Lemma 4.8.**

$$\left\|\nabla^2 b(f, d, U)\right\|_\infty \leq \frac{k(k+1)(k+2)(k+3)}{24} \left\|\nabla^2 b(f, d, \Delta)\right\|_\infty h^2.$$

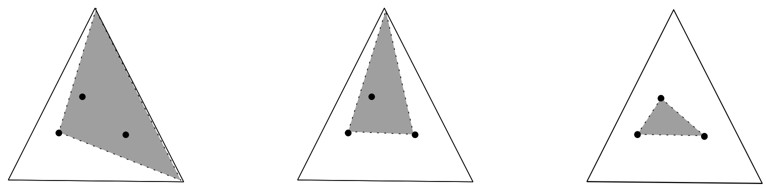


Figure 5: Reparametrization in dimension 2.

*Proof.*

- First, assume that $\left\|\nabla^2 b(f,d,\Delta)\right\|_\infty = 1$.
  Let $q_d$ be the quadratic form of the example 3.8, satisfying

  $$\nabla^2 b(q_d,d,\Delta) = (1,\ldots,1).$$

  Then $\nabla^2 b(q_d \pm f,d,\Delta) \geq 0$, from which we get $\nabla^2 b(q_d \pm f,d,U) \geq 0$: indeed, the second differences $\nabla^2 b(q_d \pm f,d,U)$ are the Bernstein coefficients of the directional derivatives of $\dfrac{-1}{d(d-1)}(q_d \pm f)$ with respect to $U$. Those coefficients can be computed by successive calls to De Casteljau's algorithm, which involve only convex sums of the (nonnegative) Bernstein coefficients of the directional derivatives of $\dfrac{-1}{d(d-1)}(q_d \pm f)$ with respect to $\Delta$.
  Hence,

  $$\left\|\nabla^2 b(f,d,U)\right\|_\infty \leq \left\|\nabla^2 b(q_d,d,U)\right\|_\infty.$$

  It is thus sufficient to study the evolution of the second differences of $q_d$ under subdivision.

- For $|\gamma| = d-2$ and $0 \leq i \leq j \leq k$, we have

  $$\nabla^2 b_{\gamma,i,j}(q_d,d,U) = \frac{-1}{d(d-1)} b\left(D_{\mathbf{u}_j-\mathbf{u}_{j-1}} D_{\mathbf{u}_i-\mathbf{u}_{i-1}} q_d, d, U\right).$$

  But

  $$D_{\mathbf{u}_j-\mathbf{u}_{j-1}} D_{\mathbf{u}_i-\mathbf{u}_{i-1}} q_d = \sum_{p=1}^{k}\sum_{q=1}^{k} (\mathbf{u}_j-\mathbf{u}_{j-1})_p (\mathbf{u}_i-\mathbf{u}_{i-1})_q \frac{\partial^2 q_d}{\partial X_p \partial X_q}$$

  and

  $$\frac{\partial^2 q_d}{\partial X_p \partial X_q} = \frac{d(d-1)}{2}\begin{cases} m_{pq} & \text{if } p<q,\\ 2m_{pp} & \text{if } p=q,\\ m_{qp} & \text{else,}\end{cases}$$

where $M$ is the matrix defined in Example 3.8.
We then obtain

$$\left|\nabla^2 b_{\gamma,i,j}(q_d, d, U)\right| \le h^2 \sum_{p=1}^{k} \sum_{q=p}^{k} m_{pq}.$$

As, by definition, $m_{pq} = p(k-q+1)$ for $1 \le p \le q \le k$, we can easily compute the sum $\sum\limits_{p=1}^{k} \sum\limits_{q=p}^{k} m_{pq}$:

$$\begin{aligned}
\sum_{p=1}^{q} \sum_{q=p}^{k} m_{pq} &= \sum_{p=1}^{k} p \sum_{q=p}^{k} (k-q+1) = \sum_{p=1}^{k} p \sum_{\ell=1}^{k-p+1} \ell \\
&= \sum_{p=1}^{k} p \frac{(k-p+1)(k-p+2)}{2} \\
&= \frac{1}{2} \left( \sum_{p=1}^{k} p^3 - (2k+3) \sum_{p=1}^{k} p^2 + (k+1)(k+2) \sum_{p=1}^{k} p \right) \\
&= \frac{k^2(k+1)^2}{8} - \frac{(2k+3)k(k+1)(2k+1)}{12} + \frac{k(k+1)^2(k+2)}{4} \\
&= \frac{k(k+1)(k+2)(k+3)}{24}.
\end{aligned}$$

Hence

$$\left|\nabla^2 b_{\gamma,i,j}(q_d, d, U)\right| \le h^2 \frac{k(k+1)(k+2)(k+3)}{24},$$

so that

$$\left\|\nabla^2 b(q_d, d, U)\right\|_\infty \le h^2 \frac{k(k+1)(k+2)(k+3)}{24}.$$

- In the general case, we finally get:

$$\begin{aligned}
\left\|\nabla^2 b(f, d, U)\right\|_\infty &\le \left\|\nabla^2 b(q_d, d, U)\right\|_\infty \left\|\nabla^2 b(f, d, \Delta)\right\|_\infty \\
&\le \frac{k(k+1)(k+2)(k+3)}{24} \left\|\nabla^2 b(f, d, \Delta)\right\|_\infty h^2,
\end{aligned}$$

as wanted. □

We can now deduce the following theorem, expressing the convergence of the control points under subdivision:

**Theorem 4.9.** *Let $\Delta = U^1 \cup \cdots \cup U^s$ be a subdivision of the standard simplex $\Delta$, and $h$ be an upper bound on the diameters of the $U^i$'s.*
*Then, for each $i \in \{1, \ldots, s\}$ and $|\alpha| = d$, we have:*

$$\left|f\left(\mathbf{v}_\alpha(d, U^i)\right) - b_\alpha(f, d, U^i)\right| \le h^2 d \frac{k^2(k+1)(k+2)^2(k+3)}{576} \left\|\nabla^2 b(f, d, \Delta)\right\|_\infty.$$

*Proof.* This is an immediate consequence of Theorem 3.12 and the previous Lemma. □

**Remark 4.10.** *Note that, by affine transformation, the result holds for every simplex $V \subset \mathbb{R}^k$ instead of $\Delta$.*

**Remark 4.11.** *Note that the previous theorem not only express the convergence of the control points under subdivision, but also shows that the rate is quadratic in the diameter of the subsimplices. This result was already known ([D]). The interest of Theorem 4.9 lies in the fact that the bound is explicit in terms of the dimension, the degree and the Bernstein expansion of $f$ over $\Delta$.*

Thus, any subdivision scheme that reduces the diameter of the subsimplices will make the control points converge to the discrete graph of $f$. We now present two such schemes.

The first one is based on successive standard triangulation of degree 2. Indeed, consider the standard triangulation $T_{2^N}(\Delta)$ $(N \geq 1)$.

**Lemma 4.12.** *If $U \in T_{2^N}(\Delta)$, and $h$ its diameter. Then:*

$$h \leq \frac{\sqrt{k}}{2^N}.$$

*Proof.* Write

$$U = [\mathbf{u}_0, \ldots, \mathbf{u}_k]$$

the vertices of the simplex $U$. Let $\mathbf{u}_i$ and $\mathbf{u}_j$ $(i < j)$ be the vertices of an edge whose length is the diameter $h$ of $U$. Consider a piecewise line curve joining the vertices $\mathbf{u}_i$,..., $\mathbf{u}_j$ along the directions $\mathbf{e}_\ell$. The length of a segment joining two consecutive vertices is $\frac{1}{2^N}$, by definition of $T_{2^N}(\Delta)$. The Pythagorean theorem then implies that the distance between $\mathbf{u}_i$ and $\mathbf{u}_j$ is less than $\frac{\sqrt{k}}{2^N}$. □

The so-called binary splitting is another subdivision scheme that reduces the diameter of the subsimplices. It consists in splitting each simplex at the midpoint of its longest edge (note that it is not necessary unique). In this case, a call to De Calsteljau algorithm can be executed faster than for an arbitrary point in the interior of the simplex ([P]).

**Lemma 4.13.** *After at most $\frac{k(k+1)}{2}$ steps of binary splitting of a simplex of diameter $h$, the diameter of the subsimplices is less than $h/2$.*

*Proof.* Assume that $V$ has been split into $U^1 \cup U^2$ by halving its longest edge. Then the edges of $U^1$ cannot be longer than the diameter $h$ of $V$, and one of them is less than $h/2$. This easily implies the result, since in each subsimplex, there are $\frac{k(k+1)}{2}$ edges. □

# 5 Certificates of positivity

Let $f \in \mathbb{R}[\mathbf{X}]$ be a multivariate polynomial of degree $d$. We aim at deciding if $f$ is positive on a simplex $V$, and in the same time obtaining a certificate of positivity, *i.e.* an algebraic identity making its positivity trivial (one can think of a "one-line proof"). We now present how the simplicial Bernstein basis can be used. Without loss of generality, we can assume that $f$ is studied on the standard simplex $\Delta$.

## 5.1 By degree elevation

The main idea is to consider $f$ as a polynomial of degree $D \geq d$. For $D$ big enough, the Bernstein coefficients of $f$ converge to values of $f$ at its grid points, and then satisfy the certificate of positivity $\mathrm{Cert}(b(f, D, V))$. The following theorem gives a bound on the degree $D$ to be considered :

**Theorem 5.1.** *Let $f \in \mathbb{R}[\mathbf{X}]$ be a polynomial of degree $d$, positive on the standard simplex $\Delta$.*
*Let $m$ be the minimum of $f$ over $\Delta$.*
*Assume that*

$$D > \frac{k(k+2)d(d-1)}{24m} \left\| \Delta^2 b(f, d, \Delta) \right\|_\infty .$$

*Then $f$ satisfies the certificate of positivity $Cert(b(f, D, \Delta))$.*

*Proof.* • Let $D$ be large enough so that

$$\max_{|\alpha|=D} \left| f\left(\mathbf{v}_\alpha(D, \Delta)\right) - b_\alpha(f, D, \Delta) \right| \leq m.$$

Then all the Bernstein coefficients of $f$ of degree $D$ are nonnegative, and the interpolation property shows that $f$ satisfies $Cert(b(f, D, \Delta))$.

• By theorem 4.2,

$$\max_{|\alpha|=D} \left| f\left(\mathbf{v}_\alpha(D, \Delta)\right) - b_\alpha(f, D, \Delta) \right| \leq \frac{k(k+2)}{24} \frac{d(d-1)}{D-1} \left\| \Delta^2 b(f, d, \Delta) \right\|_\infty ,$$

which leads to the result. □

**Remark 5.2.** *Powers et Reznick ([RP]) have also proved the following bound :*

**Theorem 5.3.** *If $D > \dfrac{d(d-1)}{2} \dfrac{\max_{|\alpha|=d} |b_\alpha(f, d, \Delta)|}{m}$, then $f$ satisfies the certificate of positivity $Cert(b(f, D, \Delta))$.*

*This bound does not depend on the dimension $k$.*

*It is worth noting that our bound is always better if $k = 1$, and that in higher dimensions, neither is better.*

## 5.2 By subdivision

The main idea is not to increase the degree anymore, but to proceed to successive standard triangulations of degree 2, in order for the Bernstein coefficients to converge to the discrete graph. This will lead to local certificates of positivity, in the following sense:

**Definition 5.4.** *Let $f$ be as above, and $S(\Delta) = \left(U^1, \ldots, U^l\right)$ be a subdivision of the simplex $\Delta$, i.e. $\Delta = U^1 \cup \cdots \cup U^l$ and the interiors of the simplices $U^i$ are disjoint.*
*If $f$ satisfies the certificates of positivity $Cert\left(b(f, d, U^i)\right)$ for all $i = 1, \ldots, l$, we say that $f$ satisfies the local certificate of positivity associated to the subdivision $S(\Delta)$, which we write $Cert\big(b\,(f, d, S(\Delta))\,\big)$.*

In the following, we will consider subdivision schemes that reduce the diameter in the following sense:

**Definition 5.5.** *Let $V$ be a nondegenerate simplex of $\mathbb{R}^k$, $S$ be a subdivision scheme (i.e. a rule for subdividing any simplex), and $S(V) = U^1 \cup \cdots \cup U^l$ be the resulting subdivision of $V$.*

- *The mesh of $S(V)$, denoted by $m(S(V))$, is the largest diameter of the subsimplices $U^i$.*
- *$S$ is said to have a shrinking factor $0 \leq C \leq 1$ if for every simplex $V$,*
$$m(S(V)) \leq Cm(V).$$
- *If $S$ is a subdivision scheme, we write $S^N(V)$ the subdivision of $V$ obtained after $N$ successive subdivision steps.*

**Example 5.6.** *The subdivision scheme consisting in $\dfrac{k(k+1)}{2}$ steps of binary splitting has a shrinking factor $\dfrac{1}{2}$ (see Lemma 4.13).*

**Theorem 5.7.** *Let $f \in \mathbb{R}[\mathbf{X}]$ be a polynomial of degree $d$, positive on $\Delta$. Let $m$ be the minimum of $f$ over $\Delta$.*
*Let $N \in \mathbb{N}^*$ be an integer and $S$ a subdivision scheme with shrinking factor $C$.*

*If $\dfrac{1}{C^N} > \dfrac{k(k+2)\sqrt{2d(k+1)(k+3)}}{24\sqrt{m}}\sqrt{\|\Delta^2 b(f, d, \Delta)\|_\infty}$, then $f$ satisfies the local certificate of positivity associated to $S^N(\Delta)$.*

*Proof.* It is sufficient to show that $\max\limits_{|\alpha|=d} |f\left(\mathbf{v}_\alpha(d, U)\right) - b_\alpha(f, d, U)| \leq m$ on each simplex $U$ of the subdivision $S^N(\Delta)$.
Theorem 4.9 implies that

$$|f\left(\mathbf{v}_\alpha(d, U)\right) - b_\alpha(f, d, U)| \leq (\sqrt{2}C^N)^2 d\frac{k^2(k+1)(k+2)^2(k+3)}{576}\left\|\nabla^2 b(f, d, \Delta)\right\|_\infty,$$

which allows us to conclude. □

## 5.3 Algorithm for computing the certificate of positivity

In order to turn the previous Theorem into a symbolic algorithm deciding if a polynomial $f \in \mathbb{Z}[\mathbf{X}]$ is positive on $\Delta$ and if so, providing a certificate of positivity, we need a stopping criterion. A lower bound on the minimum of a positive polynomial, expressed in terms of the degree, the number of variables and the bitsize of the coefficients is needed. Such a bound was first established in [BLR], and improved in [JP]:

**Theorem 5.8** ([JP])**.** *For every $f \in \mathbb{Z}[X_1, ..., X_k]$ with degree $d$ and coefficients of bitsize at most $\tau$, whose minimum $m$ over the standard simplex $\Delta$ is positive, we have*

$$m \geq 2^{-(\tau+1)d^{k+1}} d^{-(k+1)d^k} \binom{d+k}{k+1}^{-d^k(d-1)}.$$

*Taking into account that*

$$\binom{d+k}{k+1} \leq d^{k+1},$$

*the following simplified bound holds:*

$$m \geq 2^{-(\tau+1)d^{k+1}} d^{-(k+1)d^{k+1}}.$$

From what preceeds, we dispose of the following algorithm :

**Algorithm 5.9** (Local certificate of positivity)**.**

***Require:***
*-the list $b = (b_\alpha(f, d, \Delta))_{|\alpha|=d}$ of the Bernstein coefficients of a polynomial $f \in \mathbb{R}[\mathbf{X}]$ of degree $d$ with respect to the standard simplex $\Delta$ of $\mathbb{R}^k$*
*- a subdivision scheme $S$ with shrinking factor $C$.*

***Ensure:***
*- A subdivision $S^N(\Delta)$ and a local certificate of positivity associated to $S^N(\Delta)$ satisfied by $f$ if $f$ is positive on $\Delta$*
*- Else, a pair $(x, f(x))$ with $x \in \Delta$ and $f(x) < m_{k,d,\tau}$, where $m_{k,d,\tau}$ is the lower bound on the minimum obtained in theorem 5.8.*

***Initialization*** *:*
$N := 0$
$A := ((\Delta, b, n))$
$P := \emptyset$
$m_{k,d,\tau} := 2^{-(\tau+1)d^{k+1}} d^{-(k+1)d^{k+1}}$
$\nabla := \left\|\nabla^2 b(f, d, \Delta)\right\|_\infty.$

***while*** $A \neq \emptyset$ *and* $\frac{1}{C^N} \leq \frac{k(k+2)\sqrt{2d(k+1)(k+3)}}{24\sqrt{m_{k,d,\tau}}}\sqrt{\nabla}$ ***do***

*Take the first element* $(U, b, n)$ *of* $A$*, with* $U = [\mathbf{u}_0, \ldots, \mathbf{u}_k]$, $b = (b_\alpha(f, d, U))_{|\alpha|=d}$ *et* $n \geq 0$*, and remove it from* $A$*.*

***for*** $i$ *from* $0$ *to* $k$ ***do***
  ***if*** $b_{d\mathbf{e}_i}(f, d, U) \leq 0$ ***then***
    *Return* $(\mathbf{u}_i, f(\mathbf{u}_i))$*.*
  ***end if***
***end for***

***if*** $\mathrm{Cert}(b)$ ***then***
  *Add* $(U, b, n))$ *to* $P$
***else***
  *Subdivide* $U : S(U) = (U_1, \ldots, U_\ell)$*.*
  *Compute the Bernstein coefficients* $b^{(i)} = b(f, d, U_i)$ *of* $f$ *over each simplex* $U_i$*.*
  *Add the triples* $\big(U_i, b^{(i)}, n+1\big)$ *to the end of* $A$*.*
***end if***

*Take the first element* $(U, c, n)$ *of* $V$
$N := n$
***end while***

***if*** $A = \emptyset$ ***then***
  *Return* $P$*.*
***else***
  *Take the first element* $(U, b, n)$ *of* $A$*.*
  *Find* $\alpha$ *such that* $|\alpha| = d$ *and* $b_\alpha(f, d, U) < 0$*.*
  *Return* $(\mathbf{v}_\alpha(d, U), f(\mathbf{v}_\alpha(d, U)))$*.*
***end if***

## 5.4 An example

**Example 5.10.** *Let* $f = 2401x_1^4 - 1078x_1^3x_2 - 8993x_1^2x_2^2 + 2046x_1x_2^3 + 8649x_2^4 + 3822x_1^3 - 1642x_1^2x_2 - 7078x_1x_2^2 + 1488x_2^3 - 5045x_1^2 + 850x_1x_2 + 12526x_2^2 - 5226x_1 + 1072x_2 + 4490$.

*Here are the simplices generated by the algorithm to prove that* $f$ *is positive on* $\Delta$*.*

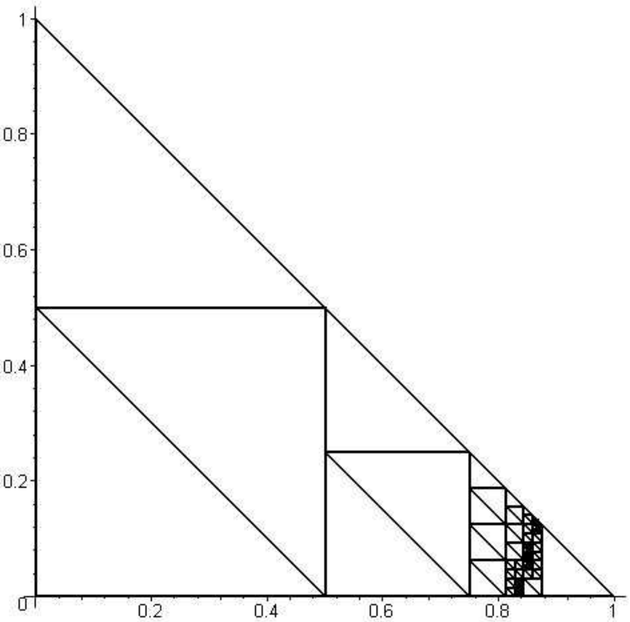



Figure 6: Successive standard triangulations of degree 2: 106 simplices

*In practice, binary splitting is usually the most efficient subdivision scheme, regarding the running time as well as the size of the certificates.*

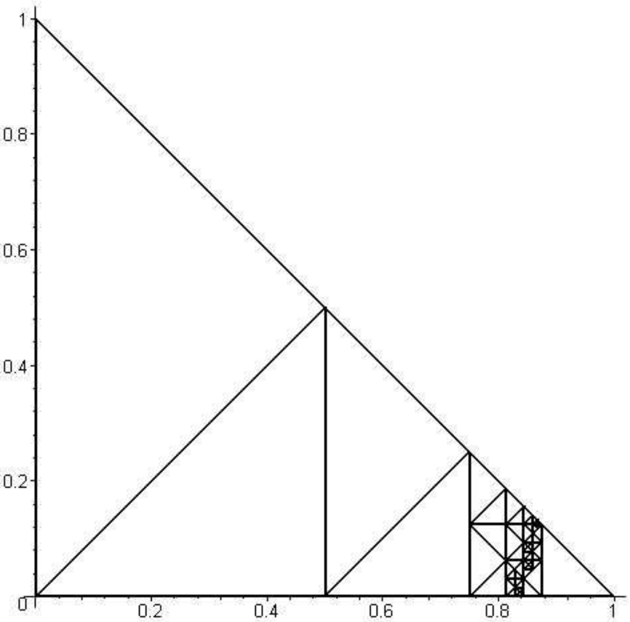



Figure 7: Successive binary splittings: 59 simplices

**Remark 5.11.** *The algorithm is adapted to the geometry of $f$, in the following sense : if $f$ is sufficiently positive to verify the certificate of positivity on a simplex $U$ occuring at some step of the algorithm, then this simplex will not be subdivided. The subdivision is only refined locally, where $f$ is small. This is a huge difference with degree elevation techniques, which are global methods.*

# References

[AG] E. Allgower, K. Georg, Introduction to Numerical Continuation Methods, SIAM Classics in Applied Mathematics 45. 2003.

[BCR] F. Boudaoud, F. Caruso, M.-F. Roy, Certificates of Positivity in the Bernstein Basis , Discrete and Computational Geometry, Volume 39, Number 4, 639-655 (2008)

[BLR] S. Basu, R. Leroy, M.-F. Roy, A bound on the minimum of a real positive polynomial over the standard simplex, preprint

[D] W. Dahmen, Subdivision algorithms converge quadratically, Journal of Computational and Applied Mathematics, Volume 16, Issue 2, 145 - 158 (1986)

[DM] W. Dahmen and C. A. Micchelli, Convexity of multivariate Bernstein polynomials and box spline surfaces, Stud. Sci. Math. Hung. 23 : 265-287 (1988)

[GP] T. Goodman, J. Peters, Bézier Nets, Convexity and Subdivision on Higher Dimensional Simplices, Computer Aided Geometric Design 12(1): 53-65 (1995)

[JP] G. Jeronimo, D. Perrucci, On the minimum of a positive polynomial over the standard simplex, preprint

[KP] L. Kobbelt, H. Prautzsch, Convergence of subdivision and degree elevation , Advances in Computational Mathematics, Volume 2, Number 1 (1994), 143-154

[L] R. Leroy, Une borne effective sur l'écart entre le polytope de contrôle et le graphe d'un polynôme réel sur un simplexe, to appear in Comptes Rendus de l'Académie des Sciences.

[NPL] D. Nairn, J. Peters, D. Lutterkort, Sharp, quantitative bounds on the distance between a polynomial piece and its Bézier control polygon, Computer Aided Geometric Design 16(7): 613-631 (1999)

[P] J. Peters, Evaluation and approximate evaluation of the multivariate Bernstein-Bézier form on a regularly partitioned simplex, ACM Transactions on Mathematical Software, Volume 20 , Issue 4 : 460 - 480 (1994)

[R] U. Reif, Best bounds on the approximation of polynomials and splines by their control structure, Computer Aided Geometric Design 17(6): 579 - 589 (2000)

[RP] B. Reznick, V. Powers, Polynomials positive on unbounded rectangles, Positive polynomials in Control (A. Garulli, D. Henrion eds.), Lecture Notes in Control and Inform. Sci. , 312 : 151-163 , Springer, Berlin (2005)